

中文界面諮詢委員會

簡介《香港增補字符集-2001》

目的

- 1 本文件旨在向各委員簡介在二零零一年十二月公布的《香港增補字符集-2001》。

背景

- 2 香港特區政府在「數碼 21」資訊科技策略下，為香港特別行政區制訂一套共通的電子中文界面，以方便市民能更有效地使用中文進行電子通訊。共通電子中文界面的基礎包括採用 ISO 10646 國際編碼標準及制訂《香港增補字符集》。
- 3 在一九九九年九月，香港特別行政區政府與「中文界面諮詢委員會」(中諮會)公布所合作編製的《香港增補字符集》。《香港增補字符集》包括 4,702 個中文字符，該等字符是供給市民及政府部門在以中文處理電子資訊時有需要使用的，而不包括在大五碼內的字符。制訂《香港增補字符集》的目的，是令市民及政府共同使用一套包含香港特有用字的字符集，以能更有效及準確地進行中文電子通訊及資料交換。
- 4 鑑於市民及政府部門不時有需要申請把新字符納入《香港增補字符集》內，政府與中諮會於二零零零年四月公布了《香港增補字符集》字符增收程序和原則。自公布該程序和原則以來及截至二零零一年十二月為止，中諮會經處理市民及政府部門提交的新增字符申請後，共增收了 116 個字符(請參閱附件一)。

《香港增補字符集-2001》

- 5 在第十二次「中文資訊科技工作小組」會議上，各組員同意在 ISO 10646 國際編碼標準的新版本(ISO/IEC 10646-2:2001)公布後，是適當的時候將《香港增補字符集》作出修訂以作配合。ISO/IEC 10646-2:2001 已於二零零一年十一月獲得公布，政府亦於二零零一年十二月三十一日發放《香港增補字符集 - 2001》。

- 6 《香港增補字符集 - 2001》較一九九九年公布的《香港增補字符集》，新增了 116 個中文字，合共收納了 4,818 個字符。其中 2,592 個字符屬於統一的中日韓表意文字字符、540 個字符屬於擴展區 A、1,640 個字符屬於擴展區 B、11 個字符屬於兼容字符區、而餘下的 35 個字符則放在 ISO 10646 國際編碼標準的私人使用區。
- 7 有兩個在《香港增補字符集》一九九九年版本內的字符錯誤對應個別的 ISO 10646 碼位。在《香港增補字符集 - 2001》的版本內已作出有關更正，詳情請參閱附件二。
- 8 我們已把《香港增補字符集 2001》的技術性文件及樣版文字輸入工具，存放在「數碼 21」資訊科技策略網站上供市民下載（http://www.digital21.gov.hk/chi/structure/struct_stand.html）為方便查閱，我們亦將有關《香港增補字符集》一九九九年版本的技術性文件繼續保留在該網站上。

資訊科技署
二零零二年六月

《香港增補字符集 - 2001》的新增字符

自二零零零年四月公布《香港增補字符集》字符增收程序和原則以來，中諮會經處理市民及政府部門提交的新增字符申請後及截至二零零一年十二月為止，共增收了 116 個字符。其中 106 個字符已獲收納在 ISO 10646 國際編碼標準的 ISO/IEC 10646-2:2001 版本內，剩下的 10 個字符則將放在其私人使用區。有關 116 個新增字符及它們在 ISO 10646 國際編碼標準內的分佈，請參閱以下的列表。

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001
1	椰	8C40	503B	503B	503B
2	傘	8C41	6DFE	6DFE	6DFE
3	鸞	8C42	F579	F579	2-9C73
4	熇	8C43	F57A	F57A	F57A
5	尉	8C44	F57B	3DC9	3DC9
6	祐	8C45	888F	888F	888F
7	灑	8C46	F57D	F57D	2-414E
8	弁	8C47	7077	7077	7077
9	岷	8C48	5CF5	5CF5	5CF5
10	霏	8C49	F580	4B20	4B20
11	睥	8C4A	F581	F581	2-51CD
12	窺	8C4B	F582	3559	3559
13	鑄	8C4C	F583	F583	2-5D30
14	愧	8C4D	6122	6122	6122
15	鋳	8C4E	F585	F585	2-8A32
16	辦	8C4F	8FA7	8FA7	8FA7
17	鉞	8C50	91F6	91F6	91F6
18	嫌	8C51	7191	7191	7191
19	𠵼	8C52	6719	6719	6719
20	壘	8C53	73BA	73BA	73BA
21	替	8C54	F58B	F58B	2-3281
22	鷓	8C55	F58C	F58C	2-A107
23	龜	8C56	F58D	3C8B	3C8B
24	𠵼	8C57	F58E	F58E	2-1980
25	颯	8C58	F58F	4B10	4B10
26	礮	8C59	78E4	78E4	78E4
27	瑄	8C5A	7402	7402	7402
28	江	8C5B	51AE	51AE	51AE

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001
29	邳	8C5C	F593	F593	2-870F
30	齋	8C5D	F594	4009	4009
31	櫛	8C5E	6A63	6A63	6A63
32	齧	8C5F	F596	F596	2-A2BA
33	鎩	8C60	F597	4223	4223
34	齧	8C61	860F	860F	860F
35	穰	8C63	7A2A	7A2A	7A2A
36	駟	8C64	F59B	F59B	2-9947
37	鑊	8C65	F59C	F59C	2-8AEA
38	靖	8C66	9755	9755	9755
39	瀾	8C67	704D	704D	704D
40	匡	8C68	5324	5324	5324
41	帛	8C69	F5A0	F5A0	2-207E
42	鎔	8C6A	93F4	93F4	93F4
43	肅	8C6B	76D9	76D9	76D9
44	鏡	8C6C	F5A3	F5A3	2-89E3
45	曙	8C6D	F5A4	F5A4	F5A4
46	矜	8C6E	77DD	77DD	77DD
47	汴	8C6F	4EA3	4EA3	4EA3
48	俶	8C70	4FF0	4FF0	4FF0
49	僕	8C71	50BC	50BC	50BC
50	丰	8C72	4E2F	4E2F	4E2F
51	众	8C73	4F17	4F17	4F17
52	隗	8C74	F5AB	F5AB	F5AB
53	吳	8C75	5434	5434	5434
54	紘	8C76	7D8B	7D8B	7D8B
55	墻	8C77	5892	5892	5892
56	爾	8C78	58D0	58D0	58D0
57	峰	8C79	F5B0	F5B0	2-1DB6
58	庄	8C7A	5E92	5E92	5E92
59	庙	8C7B	5E99	5E99	5E99
60	穰	8C7C	5FC2	5FC2	5FC2
61	愈	8C7D	F5B4	F5B4	2-2712
62	齋	8C7E	658B	658B	658B
63	杓	8CA1	F5B6	F5B6	2-33F9
64	椳	8CA2	6919	6919	6919
65	櫟	8CA3	6A43	6A43	6A43
66	氫	8CA4	F5B9	F5B9	2-3C63
67	浪	8CA5	6CFF	6CFF	6CFF
68	熾	8CA7	7200	7200	7200

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001
69	𧈧	8CA8	F5BD	F5BD	2-4505
70	玨	8CA9	738C	738C	738C
71	瑒	8CAA	F5BF	3EDB	3EDB
72	琛	8CAB	F5C0	F5C0	2-4A13
73	嬋	8CAC	5B15	5B15	5B15
74	璫	8CAD	74B9	74B9	74B9
75	讚	8CAE	8B83	8B83	8B83
76	簾	8CAF	F5C4	F5C4	2-5CA4
77	褔	8CB0	F5C5	F5C5	2-5695
78	窓	8CB1	7A93	7A93	7A93
79	簪	8CB2	7BEC	7BEC	7BEC
80	楊	8CB3	7CC3	7CC3	7CC3
81	繡	8CB4	7E6C	7E6C	7E6C
82	苧	8CB5	82F8	82F8	82F8
83	菌	8CB6	8597	8597	8597
84	斬	8CB7	F5CC	F5CC	F5CC
85	祕	8CB8	8890	8890	8890
86	赴	8CB9	F5CE	F5CE	F5CE
87	躑	8CBA	8EB9	8EB9	8EB9
88	幹	8CBB	F5D0	F5D0	F5D0
89	込	8CBC	8FCF	8FCF	8FCF
90	發	8CBD	855F	855F	855F
91	聊	8CBE	99E0	99E0	99E0
92	鍾	8CBF	9221	9221	9221
93	鋼	8CC0	F5D5	F5D5	F5D5
94	關	8CC1	F5D6	F5D6	2-8DB9
95	埴	8CC2	F5D7	F5D7	2-143F
96	曉	8CC3	F5D8	4071	4071
97	漿	8CC4	F5D9	42A2	42A2
98	媧	8CC5	5A1A	5A1A	5A1A
99	奠	8CC9	9868	9868	9868
100	杵	8CCA	676B	676B	676B
101	簍	8CCB	F5E0	4276	4276
102	坳	8CCC	573D	573D	573D
103	瞽	8CCE	85D6	85D6	85D6
104	斑	8CCF	F5E4	F5E4	2-497B
105	苈	8CD0	82BF	82BF	82BF
106	翹	8CD1	F5E6	F5E6	2-710D
107	鰓	8CD2	F5E7	4C81	4C81
108	蔡	8CD3	F5E8	F5E8	2-6D74

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001
109	嶮	8CD4	5D7B	5D7B	5D7B
110	苜	8CD5	F5EA	F5EA	2-6B15
111	藎	8CD6	F5EB	F5EB	2-6FBE
112	騏	8CD7	F5EC	F5EC	F5EC
113	驥	8CD8	F5ED	F5ED	F5ED
114	弘	8CD9	5B96	5B96	5B96
115	鉞	8CDA	F5EF	F5EF	F5EF
116	蕞	8CDC	7E5B	7E5B	7E5B

註：ISO/IEC 10646-2:2001 的出版是補充及與 ISO/IEC 10646-1:2000 並存使用。

《香港增補字符集-2001》字符的編碼更正

就《香港增補字符集》已收納的下列二個字符，已在《香港增補字符集-2001》內作出更正如下：

- 1 在大五碼字集內的“龜”字，其對應的 ISO 10646 編碼是 9F9C。《香港增補字符集》內收納了另一類似的“龜”字，其對應的 ISO 10646 編碼亦是 9F9C。當上述兩字符在 ISO 10646 國際編碼標準及大五碼間進行碼位轉換時，可能會產生錯誤的轉換結果。

《香港增補字符集》收納“龜”的字符，是基於源字集分別編碼原則 (Source Separation Rule)。

為避免上述“龜”和“龜”兩字之間可能出現的轉換錯誤問題，收納在《香港增補字符集-2001》內的“龜”字的 ISO 10646 對應編碼已由 9F9C 改為 F907。有關的編碼如下：

大五碼	在一九九九年出版的《香港增補字符集》內對應的 ISO 10646 編碼 (ISO/IEC 10646-1:1993) (ISO/IEC 10646-1:2000)	在《香港增補字符集 - 2001》內對應的 ISO 10646 編碼 (ISO/IEC 10646-1:1993) (ISO/IEC 10646-1:2000) (ISO/IEC 10646-2:2001)
C074 龜	9F9C	9F9C
8BF8 龜	9F9C	F907

- 2 在一九九九年出版的《香港增補字符集》內的“櫛”字，其對應大五碼和 ISO 10646 (ISO10646-1:2000 版本)的編碼分別是 A0E8 和 EEA1。由於在 ISO 10646-1:2000 版本內此字的編碼其實是 3C18，現已在《香港增補字符集-2001》內就該字符對應的 ISO 10646 編碼由 EEA1 更正為 3C18。